1. Các thuộc tính không ảnh hưởng:
2. Họ tên, địa chỉ, dân tộc, Sở thích không liên quan công việc
3. Lương: bị ảnh hưởng chính bởi trình độ, các yếu tố chính ảnh hưởng
4. Bằng cấp: Loại bằng, xếp loại, trường đào tạo
5. Tuổi
6. Giới tính
7. Mảng công việc
8. Số năm kinh nghiệm
9. Sở trường trong công việc: C++, NET,…
10. Độ hót của công việc hiện tại: Lap trinh
11. Ngoại ngữ: tieng anh (ielst), tieng anh toiec, tieng nhat, tieng trung,…
12. Vị trí ứng tuyển: Teamlead
13. Hệ số trách nhiệm
14. Giải thưởng, thành tích cá nhân.
15. Ngưỡng lương: 7tr- 20tr
16. Xếp loại trình độ (năng lực công việc) : A, B, C, D (thuộc tính này cho sẵn. Nếu tính thì có thể dựa vào lịch sử kết quả công việc đã đảm nhận)
17. Ý thức, trách nhiệm với công việc, công ty
18. Thời gian gắn bó với công ty
19. Mức lương công ty cũ trả
20. Lương (Nhãn)

Note: Các thuộc tính cần động, tích lũy hoặc thay đổi theo thời gian

Note: Với dạng bài toán này nếu tính toán hoặc quy về bài toán tìm kiếm so sánh thì cần vector hóa. VD: regression, tìm kiếm. Nếu để các trường dạng phi số thì không tính toán được, phải dùng các phương pháp như decision tree để xử lý.

1. Thuộc tính ảnh hưởng rating công việc

1.Thuộc tính :

- Bằng cấp

- Tuổi

- Giới tính

- Chuyên môn (Công việc đảm nhận)

- Số năm kinh nghiêm

- sở trường

- ngoại ngữ

- Lương.

Là các thuộc tính chính để đánh giá Ratings của người đố với công việc, còn thuộc tính phụ thuộc ko nhiều , ảnh hưởng ít nên ta bỏ.

2. Tính Ratings hướng 2 cách :

- Thiết lập chỉ định :

ví dụ :

1<=sonamkn<2 -> \*

2<=sonamkn<3 ->\*\*

3<=sonamkn<4 ->\*\*\*

4<=sonamkn<5 ->\*\*\*\*

sonamkn=>5 -> \*\*\*\*\*

luong :

3T-5M => 1\*

5T-7T => 2\*

7-9 =>3\*

9-11=>4\*

>=11=> 5\*

Quy định : luong chiếm chiếm 0.7 , sonamkn chiếm 0.3

Ví dụ người có 2,5 năm kn và lương 6T => 2x0.3 + 2x0.7=2 =>Ratings=2

- Dùng Classification để phân lớp từ 1 -> 5 cũng là Ratings của người đó và hướng thuật toán Kmeans,SVM,Cây quyết định ...

1. Các việc cần làm tiếp theo
2. Tạo csdl theo yêu cầu, với các thuộc tính mà một người có thể bao gồm nhiều thuộc tính con thì cần tạo bảng con. VD : thuộc tính ‘’sở trường ‘’ một người vừa có thể code tốt C# lẫn C++
3. Fake data cho các trường dữ liệu trong csdl, ràng buộc dữ liệu sao cho hợp lý
4. Cài đặt các thuật toán với bài toán và dữ liệu đã tạo để xem thử kết quả, đánh giá